

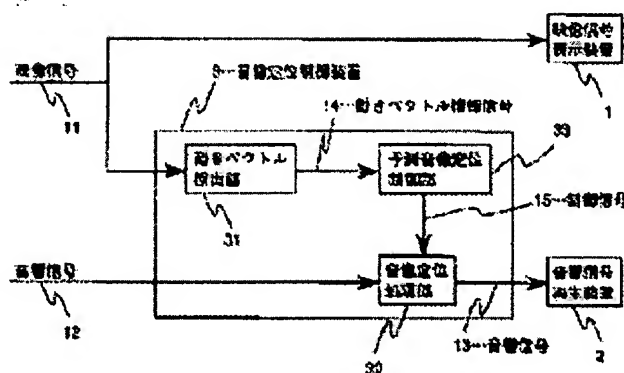
SOUND IMAGE LOCALIZATION METHOD AND SOUND IMAGE LOCALIZATION CONTROLLER AND VIDEO CONFERENCE SYSTEM

Patent number: JP8181962
Publication date: 1996-07-12
Inventor: SUZUKI TOSHIAKI; MIMURA ITARU; KAMEYAMA TATSUYA
Applicant: HITACHI LTD
Classification:
- international: **H04N7/15; H04S7/00; H04N7/15; H04S7/00; (IPC1-7): H04N7/15; H04S7/00**
- european:
Application number: JP19940320169 19941222
Priority number(s): JP19940320169 19941222

Report a data error here

Abstract of JP8181962

PURPOSE: To make a display position of a sounder matching a localized position of a sound signal from the sounder in an excellent way. **CONSTITUTION:** A motion vector detection section 31 detects a motion vector of a sounder included in a video signal 11 and provides an output of a motion vector information signal 14. A prediction sound image localization control section 33 predicts a position of the sounder at a lapse of the processing time from the detection of a motion vector of the sounder till the end of sound image localization processing and gives a control section 15 to a sound signal processing section 32. The sound signal control section 32 applies sound image localization processing (control of balance of sound reproduced from each speaker and reproduction timing or the like) to the sound signal 12 based on the control signal 15 and provides an output of a sound signal 13 after the processing. Since the sound image is localized at a position that takes a time required for the sound image localization processing into account, the display position of the sounder matches a localized position of the sound signal from the sounder in an excellent way. Thus, the reproduction of sound with full of presence is attained.



Data supplied from the esp@cenet database - Worldwide

" "

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-181962

(43) 公開日 平成8年(1996)7月12日

(51) Int.Cl.⁶

識別記号

庁内整理番号

F I

技術表示箇所

H 0 4 N 7/15

H 0 4 S 7/00

F

審査請求 未請求 請求項の数 8 O L (全 10 頁)

(21) 出願番号 特願平6-320169

(22) 出願日 平成6年(1994)12月22日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 鈴木 敏明

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 三村 到

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 亀山 達也

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(74) 代理人 弁理士 有近 紳志郎

(54) 【発明の名称】 音像定位方法および音像定位制御装置およびテレビ会議システム

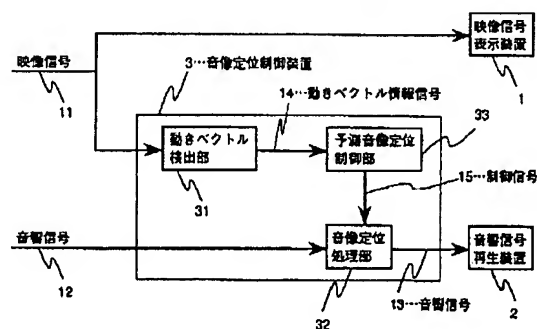
(57) 【要約】

【目的】 発音体の表示位置と発音体に付随した音響信号の定位位置とを良好に一致させる。

【構成】 動きベクトル検出部31は、映像信号11に含まれる発音体の動きベクトルを検出し、動きベクトル情報信号14を出力する。予測音像定位制御部33は、発音体の動きベクトル検出から音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測し、制御信号15を音響信号処理部32に入力する。音響信号制御部32は、前記制御信号15に基づいて、音響信号12に対して音像定位処理（各スピーカから再生する音量バランスおよび再生タイミング等を制御する）を行い、処理後の音響信号13を出力する。

【効果】 音像定位処理に要する時間を考慮した位置に音像定位するため、発音体の表示位置と音像定位位置とを良好に一致させることが出来る。これにより、臨場感ある音響再生が可能となる。

(図1)



1

【特許請求の範囲】

【請求項1】 映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出し、発音体の動きベクトル検出から音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測し、音響信号に対して前記予測した位置への音像定位を行うことを特徴とする音像定位方法。

【請求項2】 映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出し、音響信号に対して前記動きベクトルに基づく音像定位を行い、前記動きベクトルの検出から音像定位処理の終了までの処理時間だけ前記映像信号の表示を遅延させることを特徴とする音像定位方法。

【請求項3】 映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出する動きベクトル検出手段と、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測する予測音像定位制御手段と、音響信号に対して前記予測した位置への音像定位を行う音像定位処理手段とを具備したことを特徴とする音像定位制御装置。

【請求項4】 映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出する動きベクトル検出手段と、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測する予測音像定位制御手段と、ある予測位置と次の予測位置の間に補間により複数の補間位置を設定する補間音像定位制御手段と、音響信号に対して前記補間位置および予測位置への音像定位を行う音像定位処理手段とを具備したことを特徴とする音像定位制御装置。

【請求項5】 請求項3または請求項4に記載の音像定位制御装置において、前記予測音像定位制御手段は、前記音像定位処理の終了までの処理時間を予測ベクトル算出時間 c とし、フレーム間時間を t とし、動きベクトルにおける単位時間を T とすると、 $c < n \cdot t$ となる最小の正の整数 n を求め、動きベクトル $\{r \rightarrow\}$ に $n \cdot t / T$ を乗算して予測ベクトル $\{x \rightarrow\}$ を算出することを特徴とする音像定位制御装置。

【請求項6】 映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出する動きベクトル検出手段と、音響信号に対して前記動きベクトルに基づく音像定位を行う音像定位処理手段と、映像信号を一時的に蓄積し前記動きベクトルの検出から音像定位処理の終了までの処理時間だけ遅延させて前記映像信号を出力する映像蓄積処理手段とを具備したことを特徴とする音像定位制御装置。

2

【請求項7】 請求項3から請求項6のいずれかに記載の音像定位制御装置において、前記動きベクトル検出手段は、ある発音体の動きベクトルとして複数の候補があるときそれらを出力するものであり、その動きベクトル検出手段の後に、個々の発音体とそれに付随する音響信号の対応関係を維持するように前記動きベクトルの複数の候補中から1つを選定する追従音像定位制御手段を設けたことを特徴とする音像定位制御装置。

【請求項8】 映像信号および音響信号を受信する受信装置と、前記受信した映像信号に基づいて画面に画像を表示する映像信号表示手段と、前記受信した音響信号に対して音像定位処理を行う請求項3から請求項7のいずれかに記載の音像定位制御装置と、その音像定位制御装置から出力される音響信号に基づいて音響を再生する音響信号再生装置とを備えたことを特徴とするテレビ会議システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、音像定位方法および音像定位制御装置およびテレビ会議システムに関する。さらに詳しくは、発音体の画面上の位置と音像定位位置とを良好に一致させる音像定位方法および音像定位制御装置およびテレビ会議システムに関する。

【0002】

【従来の技術】従来の技術として、特開平1-296900号公報に開示の音像定位制御方式が知られている。この音像定位制御方式は、映像信号および音響信号を受信し、映像信号に含まれる顔、口、楽器等の発音体の画像パターンを識別抽出し、その画像パターンの画面上の位置に追従して前記音響信号を再生する際の音像定位を行う方式である。

【0003】

【発明が解決しようとする課題】上記従来の音像定位制御方式には、次の問題点がある。

(1) 発音体の動きを検出し追従させて音像定位するとき、処理（例えば画像フレーム間の差分演算など）による遅延が生じる。しかし、この遅延があると、発音体の画像パターンの画面上の位置と音像定位位置とが一致しなくなる。

(2) 発音体の画像パターンの画面上での移動速度が速い場合、ある音像定位位置から次の音像定位位置までの移動が不自然に飛ぶようになる。

(3) 複数の発音体の画像パターンがある場合、発音体とそれに付随する音響信号の対応関係が不明確になる。

【0004】そこで、本発明の第1の目的は、発音体の画像パターンの画面上の位置と音像定位位置とを良好に一致させる音像定位方法および音像定位制御装置およびテレビ会議システムを提供することにある。また、本発明の第2の目的は、音像定位位置を連続的に移動させる音像定位方法及び音像定位制御装置及びテレビ会議

3

システムを提供することにある。さらに、本発明の第3の目的は、複数の発音体がある場合、発音体とそれに付随した音響信号とを良好に対応させる音像定位方法および音像定位制御装置およびテレビ会議システムを提供することにある。

【0005】

【課題を解決するための手段】第1の観点では、この発明は、映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出し、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測し、音響信号に対して前記予測した位置への音像定位を行うことを特徴とする音像定位方法を提供する。

【0006】第2の観点では、この発明は、映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出し、音響信号に対して前記動きベクトルに基づく音像定位を行い、前記動きベクトルの検出から音像定位処理の終了までの処理時間だけ前記映像信号の表示を遅延させることを特徴とする音像定位方法を提供する。

【0007】第3の観点では、この発明は、映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出する動きベクトル検出手段と、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測する予測音像定位制御手段と、音響信号に対して前記予測した位置への音像定位を行う音像定位処理手段とを具備したことを特徴とする音像定位制御装置を提供する。

【0008】第4の観点では、この発明は、映像信号から画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出する動きベクトル検出手段と、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測する予測音像定位制御手段と、ある予測位置と次の予測位置の間に補間により複数の補間位置を設定する補間音像定位制御手段と、音響信号に対して前記補間位置および予測位置への音像定位を行う音像定位処理手段とを具備したことを特徴とする音像定位制御装置を提供する。

【0009】第5の観点では、この発明は、上記構成の音像定位制御装置において、前記予測音像定位制御手段は、前記音像定位処理の終了までの処理時間を予測ベクトル算出時間 c とし、フレーム間時間を t とし、動きベクトルにおける単位時間を T とすると、 $c < n \cdot t$ となる最小の正の整数 n を求め、動きベクトル $\{r \rightarrow\}$ に $n \cdot t / T$ を乗算して予測ベクトル $\{x \rightarrow\}$ を算出することを特徴とする音像定位制御装置を提供する。

【0010】第6の観点では、この発明は、映像信号か

4

ら画面上の発音体の動きベクトル（単位時間での発音体の画面上での移動方向および移動距離を表すベクトル）を検出する動きベクトル検出手段と、音響信号に対して前記動きベクトルに基づく音像定位を行う音像定位処理手段と、映像信号を一時的に蓄積し前記動きベクトルの検出から音像定位処理の終了までの処理時間だけ遅延させて前記映像信号を出力する映像蓄積処理手段とを具備したことを特徴とする音像定位制御装置を提供する。

【0011】第7の観点では、この発明は、上記構成の音像定位制御装置において、前記動きベクトル検出手段は、ある発音体の動きベクトルとして複数の候補があるときそれらを出力するものであり、その動きベクトル検出手段の後に、個々の発音体とそれに付随する音響信号の対応関係を維持するように前記動きベクトルの複数の候補中から1つを選定する追従音像定位制御手段を設けたことを特徴とする音像定位制御装置を提供する。

【0012】第8の観点では、この発明は、映像信号および音響信号を受信する受信装置と、前記受信した映像信号に基づいて画面に画像を表示する映像信号表示手段と、前記受信した音響信号に対して音像定位処理を行う上記構成の音像定位制御装置と、その音像定位制御装置から出力される音響信号に基づいて音響を再生する音響信号再生装置とを備えたことを特徴とするテレビ会議システムを提供する。

【0013】

【作用】上記第1の観点による音像定位方法および上記第3の観点による音像定位制御装置では、映像信号から画面上の発音体の動きベクトルを検出する。次に、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測する。そして、音響信号に対して前記予測した位置への音像定位を行う。このように処理の遅延を考慮した位置に音像定位するので、発音体の画面上の位置と音像定位位置とを良好に一致させることができる。

【0014】上記第2の観点による音像定位方法および上記第6の観点による音像定位制御装置では、映像信号から画面上の発音体の動きベクトルを検出する。次に、音響信号に対して前記動きベクトルに基づく音像定位を行う。さらに、前記動きベクトルの検出から音像定位処理の終了までの処理時間だけ前記映像信号の表示を遅延させる。このように処理の遅延を考慮して映像信号の表示を遅延させるので、発音体の画面上の位置と音像定位位置とを良好に一致させることができる。

【0015】上記第4の観点による音像定位制御装置では、映像信号から画面上の発音体の動きベクトルを検出する。次に、音像定位処理の終了までの処理時間を経過した時の発音体の位置を前記動きベクトルから予測する。次に、ある予測位置と次の予測位置の間に補間により複数の補間位置を設定する。そして、音響信号に対して前記補間位置および予測位置への音像定位を行う。こ

5

のように処理の遅延を考慮した位置に音像定位するので、発音体の画面上の位置と音像定位位置とを良好に一致させることが出来る。さらに、予測位置と次の予測位置との間を補間して音像定位置移動を行うため、滑らかな音像移動が可能となる。

【0016】上記第5の観点による音像定位制御装置では、予測音像定位制御手段は、前記音像定位処理の終了までの処理時間を予測ベクトル算出時間 c とし、フレーム間時間を t とし、動きベクトル $\{r \rightarrow\}$ が単位時間 T での発音体の画面上での移動方向および移動距離を表すとき、 $c < n \cdot t$ となる最小の正の整数 n を求める。次に、動きベクトル $\{r \rightarrow\}$ に $n \cdot t / T$ を乗算し、予測ベクトル $\{x \rightarrow\}$ を求める。これにより、フレームにおける発音体の位置を予測することが出来る。

【0017】上記第7の観点による音像定位制御装置では、動きベクトル検出手段は、ある発音体の動きベクトルとして複数の候補があるとき、それらを全て出力する。また、その動きベクトル検出手段の後に追従音像定位制御手段を設け、その追従音像定位制御手段は、個々の発音体とそれに付随する音響信号の対応関係を維持するように前記動きベクトルの複数の候補中から1つを選定する。これにより、複数の発音体がある場合でも、発音体とそれに付随した音響信号とを良好に対応させることが出来る。

【0018】上記第8の観点によるテレビ会議システムでは、上記構成の音像定位制御装置を備えているため、発音体の表示位置と音像定位位置の一致性を高めることが出来る。また、滑らかな音像移動が可能となる。さらに、複数の発音体がある場合でも、発音体とそれに付随する音響信号との対応関係を一致させることが出来る。

【0019】

【実施例】以下、本発明の実施例を図面を用いて説明する。なお、これにより本発明が限定されるものではない。

【0020】-第1実施例-

第1実施例は、フレームにおける発音体の位置を予測し、その予測位置に音像定位する実施例である。図9は、本発明の一実施例のテレビ会議システムのブロック図である。このテレビ会議システム100において、1は映像信号表示装置を表し、2は音響信号再生装置を表す。3は、本発明にかかる音像定位制御装置を表す。4はカメラを表し、5はマイクを表す。6は映像信号および音響信号の送受信装置を表す。また、11は映像信号を表し、12は音響信号を表す。また、13は通信回線を表す。人物(乙)は会議参加者である。

【0021】次に、各部の動作を、通信回線13によって接続された遠隔の図示せぬテレビ会議システムとのテレビ会議を想定して説明する。なお、図示せぬテレビ会議システムは、テレビ会議システム100と同じ構成であり、その会議参加者は人物(甲)である。前記カメラ

6

4は、人物(乙)を撮影し、映像信号を送受信装置6に渡す。前記マイク5は、人物(乙)の音声を受音し、音響信号を送受信装置6に渡す。前記送受信装置6は、前記カメラ4から渡された映像信号および前記マイク5から渡された音響信号を、通信回線13を介して、図示せぬテレビ会議システムへ送信する。また、前記送受信装置6は、図示せぬテレビ会議システムから通信回線13を介して送信されてきた映像信号および音響信号を受信し、映像信号11は映像信号表示装置1および音像定位制御装置3に入力し、音響信号12は音像定位制御装置3に入力する。前記映像信号表示装置1は、入力された映像信号11により画面に画像を実時間で表示する。前記音像定位制御装置3は、入力された映像信号11を解析し、その解析結果に応じて映像信号表示装置1の画面上での音像定位位置を定め、前記音響信号12に対して定位処理を行い、定位処理後の音響信号13を音響信号再生装置2に渡す。前記音響信号再生装置2は、音響信号13に基づき、音響を実時間で再生する。

【0022】図1は、前記音像定位制御装置3の構成を示すブロック図である。この音像定位制御装置3は、動きベクトル検出部31と、予測音像定位制御部33と、音像定位処理部32とから構成される。前記動きベクトル検出部31は、映像信号11から発音体の動きベクトル(単位時間 T での発音体の画面上での移動方向および移動距離を表すベクトル)を検出し、動きベクトル情報信号14を予測音像定位制御部33に入力する。なお、動きベクトルの検出は、例えば「認識技術を応用した対話型映像編集の提案(上田 他)電子情報通信学会論文誌 D-II Vol. J75-D-II No. 2 pp. 216-225 1992年2月」や「ダイナミックシーンの理解(浅田稔)電子情報通信学会編」に記載されている技術を用いることが出来る。前記予測音像定位制御部33は、入力された動きベクトル情報信号14から音響再生時の発音体の画面上の位置を予測し、その予測位置に音像定位するように指示する制御信号15を音像定位処理部32に入力する。前記制御信号15は、例えば表示位置座標を含んでいる。前記音像定位処理部32は、入力された制御信号15に基づき、音響信号再生装置2で再生する音響の音量バランスおよび再生タイミング等を制御する音像定位処理を行い、定位処理後の音響信号13を音響信号再生装置2に渡す。

【0023】図2は、第1実施例にかかる音像定位方法の説明図である。予測音像定位制御部33は、予測処理に要する時間を予測ベクトル算出時間 c とし、フレーム間時間を t (例えば $1/30$ 秒)とすると、 $c < n \cdot t$ となる最小の正の整数 n を求め、動きベクトル $\{r \rightarrow\}$ に $n \cdot t / T$ を乗算したものを予測ベクトル $\{x \rightarrow\}$ とする。この予測ベクトル $\{x \rightarrow\}$ は、現フレームから数えて n 番後のフレームまでの発音体の画面上での移動方向および移動距離を表しており、これを音像定位

7

処理部32に通知する。ここで、 n 個の予測音像定位制御部33が並列に処理する構成とすれば、フレーム毎に予測ベクトル $\{x\rightarrow\}$ を算出することが出来る。音像定位処理部32は、現在の音像定位位置の次に、現在の音像定位位置に予測ベクトル $\{x\rightarrow\}$ を加えた予測位置に音像定位する。なお、図2で、21~24は、時刻 $(k-1)$ のフレーム~時刻 (n) のフレームの画像である。91は、フレーム間時間 t である。92は、予測ベクトル算出時間 c である。合成図25は、時刻 $(k-1)$ から時刻 (n) までの発音体(甲)の表示映像を合成した図である。41は、時刻 $(k-1)$ のフレームでの表示映像である。42は、時刻 (k) のフレームでの表示映像である。43は、時刻 (n) のフレームでの表示映像である。

【0024】以上の第1実施例によれば、処理の遅延を考慮して音像定位するので、発音体の画面上の位置と音像定位位置とを良好に一致させることが出来る。また、遅延の累積もないので、発音体の画面上の位置と音像定位位置が大きくずれてしまうこともない。

【0025】なお、上記第1実施例では、動きベクトル $\{r\rightarrow\}$ として速度ベクトルを想定し、予測ベクトル $\{x\rightarrow\}$ を線形的に予測した。しかし、これに限定されず、例えば動きベクトル $\{r\rightarrow\}$ として加速度ベクトルを想定し、予測ベクトル $\{x\rightarrow\}$ を非線形的に予測してもよい。

【0026】-第2実施例-

第2実施例は、フレームにおける発音体の位置を予測し、現在の発音体の位置から予測位置の間の発音体の位置を補間し、その補間位置をたどって予測位置まで順に音像定位する実施例である。図3は、第2実施例における音像定位制御装置3Aのブロック図である。この音像定位制御装置3Aは、図1の音像定位制御装置3における予測音像定位制御部33と音像定位処理部32の間に、補間音像定位制御部34を介した構成である。前記補間音像定位制御部34では、予測音像定位制御部33から制御信号15により入力された予測位置 $(n-1)$ と予測位置 (n) の間に、補間により複数の補間位置を設定し、それら補間位置および前記予測位置を制御信号16として音響信号処理部32に入力する。

【0027】図4は、第2実施例にかかる補間音像定位方法の説明図である。合成図26は、予測音像定位制御部33において予測した発音体の予測位置 $(n-1)$ と、次の予測位置 (n) を合成した図である。補間音像定位制御部34は、予測位置 $(n-1)$ と予測位置 (n) の間を m 分割し、補間位置61~63を設定する。音響信号処理部32は、予測位置 $(n-1)$ の次に補間位置61に定位(1)し、順に補間位置62に定位(2)し、順に補間位置63まで定位 $(m-1)$ し、最後に予測位置 (n) に定位(64)する。

【0028】以上の第2実施例によれば、音像定位位置

8

を連続的に移動させることが出来る。この結果、発音体の画面上での移動速度が速い場合でも、音像定位位置が不自然に飛ばず、滑らかに移動するようになる。

【0029】-第3実施例-

第3実施例は、複数の発音体とそれらに付随した音響信号の対応を維持する実施例である。図5は、第3実施例における音像定位制御装置3Bのブロック図である。この音像定位制御装置3Bは、図3の音像定位制御装置3Aにおける動きベクトル検出部31と予測音像定位制御部33の間に、追従音像定位制御部35を介した構成である。前記動きベクトル検出部31は、ある発音体の動きベクトルとして複数の候補があるとき、それらを全て出力する。前記追従音像定位制御部35は、前フレームにおいて検出した動きベクトルと現フレームにおいて検出した動きベクトルとを比較することにより、前記複数の候補中から1つの動くベクトルを選定し、これにより個々の発音体とそれに付随する音響信号の対応関係を維持し、個々の発音体ごとに区別した動きベクトル情報信号17を予測音像定位制御部33に入力する。

【0030】図6は、第3実施例にかかる追従音像定位方法の説明図である。合成図27は、フレーム $(k-2)$ からフレーム (k) までの発音体(甲)および発音体(乙)の動きを合成した図である。説明の都合上、フレーム $(k-2)$ 、フレーム $(k-1)$ における発音体(甲)の位置40、41およびその音響信号の対応が判っているとする。また、フレーム $(k-2)$ 、フレーム $(k-1)$ における発音体(乙)の位置45、46およびその音響信号の対応が判っているとする。次のフレーム (k) では、2つの発音体の位置42、47は判るが、発音体(甲)(乙)の位置がどちらかは判らない。そこで、動きベクトル検出部31は、フレーム $(k-1)$ における発音体(甲)の位置41からフレーム (k) における2つの発音体の位置42、47への2つの動きベクトル $\{v\rightarrow\}$ (k)56および動きベクトル $\{v'\rightarrow\}$ (k)57を候補として追従音像定位制御部35に入力する。追従音像定位制御部35は、フレーム $(k-2)$ からフレーム $(k-1)$ までの発音体(甲)の動きベクトル $\{v\rightarrow\}$ ($k-1$)55となす角度 θ および θ' を算出し、角度の小さい方を発音体(甲)の動きベクトルとして選定する(角度が同じときは動きベクトルの小さい方を選定する)。これにより、位置42がフレーム (k) における発音体(甲)の位置と判る。また、位置47がフレーム (k) における発音体(乙)の位置と判る。従って、フレーム (k) における発音体(甲)の位置42およびその音響信号の対応が判る。また、フレーム (k) における発音体(乙)の位置47およびその音響信号の対応が判る。以下、同様にして、フレーム $(k+1)$ 以降でも発音体とその音響信号の対応を維持することが出来る。

【0031】図7は、第3実施例にかかる別の追従音像

定位方法の説明図である。206は、フレーム(f)での発音体(A)211および発音体(B)214の画面上の位置を表している。また、207は、フレーム(f+1)での発音体(A)212および発音体(B)215の画面上の位置を表している。次に、208および209は、フレーム(f)およびフレーム(f+1)での発音体(A)、(B)の位置を合成表示した図である。

【0032】説明の都合上、フレーム(f)における発音体(A)の位置211およびその音響信号の対応が判っているものとする。また、フレーム(f)における発音体(B)の位置214およびその音響信号の対応が判っているものとする。次のフレーム(f+1)では、2つの発音体の位置212、215は判るが、発音体(A)、(B)の位置がどちらかは判らない。つまり、208に示したように、フレーム(f)における発音体(A)211がフレーム(f+1)において位置212に移動し、且つ、フレーム(f)の発音体(B)214がフレーム(f+1)において位置215に移動する第1の移動パターンと、209に示したように、フレーム(f)における発音体(A)211がフレーム(f+1)において位置215に移動し、且つ、フレーム(f)の発音体(B)214がフレーム(f+1)において位置212に移動する第2の移動パターンとがあり、いずれであるかが判らない。

【0033】そこで、動きベクトル検出部31は、第1の移動パターンと仮定して、208における発音体(A)および(B)の動きベクトル221{a→}および動きベクトル223{b→}を算出し、発音体(A)、(B)の第1の移動候補として追従音像定位制御部35に入力する。また、第2の移動パターンと仮定して、209における発音体(A)および(B)の動きベクトル222{a'→}および動きベクトル224{b'→}を算出し、発音体A、Bの第2の移動候補として追従音像定位制御部35に入力する。追従音像定位制御部35は、フレーム(f)における発音体211を動きベクトル221{a→}だけ移動させ、フレーム(f+1)における発音体212と差分をとる。また、フレーム(f)における発音体214を動きベクトル223{b→}だけ移動させ、フレーム(f+1)における発音体215と差分をとる。そして、両差分を合計する。次に、フレーム(f)における発音体211を動きベクトル222{a'→}だけ移動させ、フレーム(f+1)における発音体215と差分をとる。また、フレーム(f)における発音体214を動きベクトル224{b'→}だけ移動させ、フレーム(f+1)における発音体212と差分をとる。そして、両差分を合計する。さらに、前記差分の合計同士を比較し、合計の小さい移動パターンが正しいと判定する。これにより、フレーム(f+1)における発音体(A)、(B)の位置が判る。また、フレーム(f+1)における発音体

(A)、(B)と音響信号の対応が判る。以下、同様に、フレーム(f+2)以降でも発音体とその音響信号の対応を維持することが出来る。

【0034】なお、上記の2通りの追従音像定位方法の外に、例えば発音体の色や輝度(R, G, B, Y, C)などの画像的特徴を比較することにより、発音体の移動先を判定する方法を用いてもよい。

【0035】以上の第3実施例によれば、複数の発音体がある場合でも、発音体とそれに付随した音響信号とを良好に対応させることが出来る。

【0036】-第4実施例-

第4実施例は、発音体の動きベクトル検出から音像定位に要する時間だけ映像信号を遅らせることにより、発音体の画面上の位置と音像定位位置とを一致させる実施例である。図8は、第4実施例における音像定位制御装置3Cのブロック図である。この音像定位制御装置3Cは、動きベクトル検出部31と、音像定位処理部37と、映像蓄積処理部36とから構成される。前記動きベクトル検出部31は、映像信号11から発音体の動きベクトルを検出し、動きベクトル情報信号18を音像定位処理部37に入力する。前記音像定位処理部37は、入力された音響信号12を前記動きベクトルにより音像定位処理し、処理後の音響信号13を音響信号再生装置2へ出力する。また、音響信号13の再生と同期して映像を表示するためのタイミングを制御する制御信号19を映像蓄積処理部36へ送る。前記映像蓄積処理部36は、映像信号11を一時的に蓄積し、前記制御信号19に基づくタイミングで取り出した映像信号11'を映像信号表示装置1へ出力する。なお、図8の動きベクトル検出部31と音像定位処理部37の間に、図5の追従音像定位制御部35および/または補間音像定位制御部34を介してもよい。

【0037】以上の第4実施例によれば、処理の遅延を考慮して映像信号の表示を遅延させるので、発音体の画面上の位置と音像定位位置とを良好に一致させることが出来る。この第4実施例は、図9のテレビ会議システム100に適用することも可能であるが、リアルタイム性があまり重要でないビデオプレー等々に適用する方が有用である。

【0038】

【発明の効果】本発明の音像定位方法および音像定位制御装置およびテレビ会議システムによれば、次の効果が得られる。

(1) 映像信号に含まれる発音体の動きを検出し、その動き検出処理および音像定位処理に要する時間後の発音体の表示位置を予測し、音響信号の定位処理を行うため、発音体の表示位置と音像定位位置との一致性を高めた音響再生が可能となる。

(2) 予測位置と次の予測位置との間を補間して音像定位移動を行うため、滑らかな音像移動が可能となる。

11

(3) 発音体が複数ある場合、発音体とそれに付属した音響信号の対応関係を常に維持するよう制御するため、発音体とそれに付属した音響信号の対応関係が明確になる。

(4) 発音体の動きを検出し、発音体の移動後の位置に音像定位するよう信号処理を行い、音響信号処理に要した時間量だけ映像信号の再生を遅延させ、映像表示および音響信号の再生を行うため、発音体の表示位置と音像定位位置との一致性を高めた音響再生が可能となる。

【図面の簡単な説明】

【図1】本発明の第1実施例にかかる音像定位制御装置を示すブロック図である。

【図2】第1実施例にかかる音像定位方法の説明図である。

【図3】本発明の第2実施例にかかる音像定位制御装置を示すブロック図である。

【図4】第2実施例にかかる補間音像定位方法の説明図である。

【図5】本発明の第3実施例にかかる音像定位制御装置

12

を示すブロック図である。

【図6】第3実施例にかかる追従音像定位方法の説明図である。

【図7】第3実施例にかかる追従音像定位方法の別の説明図である。

【図8】本発明の第4実施例にかかる音像定位制御装置を示すブロック図である。

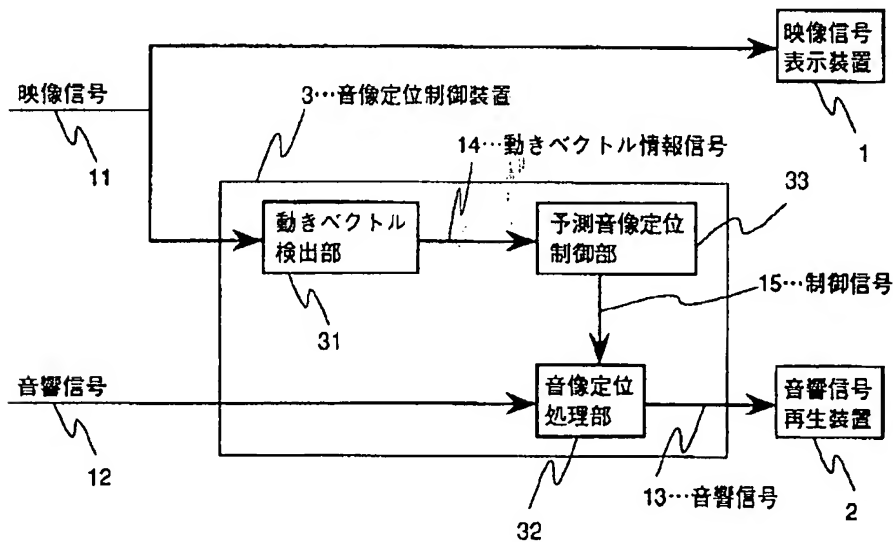
【図9】本発明の一実施例にかかるテレビ会議システムを示すブロック図である。

10 【符号の説明】

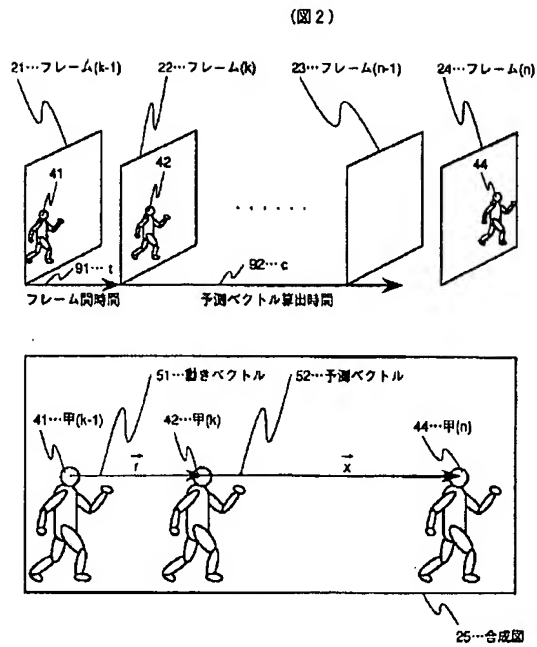
1…映像信号表示装置、2…音響信号再生装置、3、3A、3B、3C…音像定位制御装置、11、11'…映像信号、12、13…音響信号、14、18…動きベクトル情報信号、15、16、17…制御信号、31…動きベクトル検出部、32…音響信号処理部、33…予測音像定位制御部、34…補間音像定位制御部、35…追従音像定位制御部、36…映像蓄積処理部、51…動きベクトル、52…予測ベクトル

【図1】

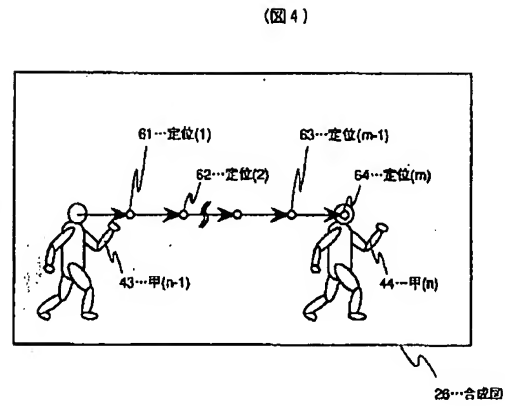
(図1)



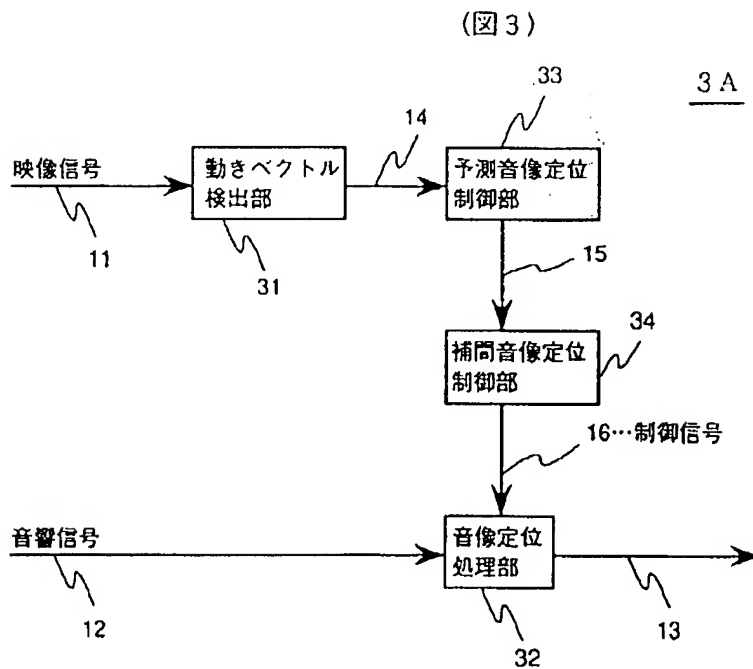
【図2】



【図4】

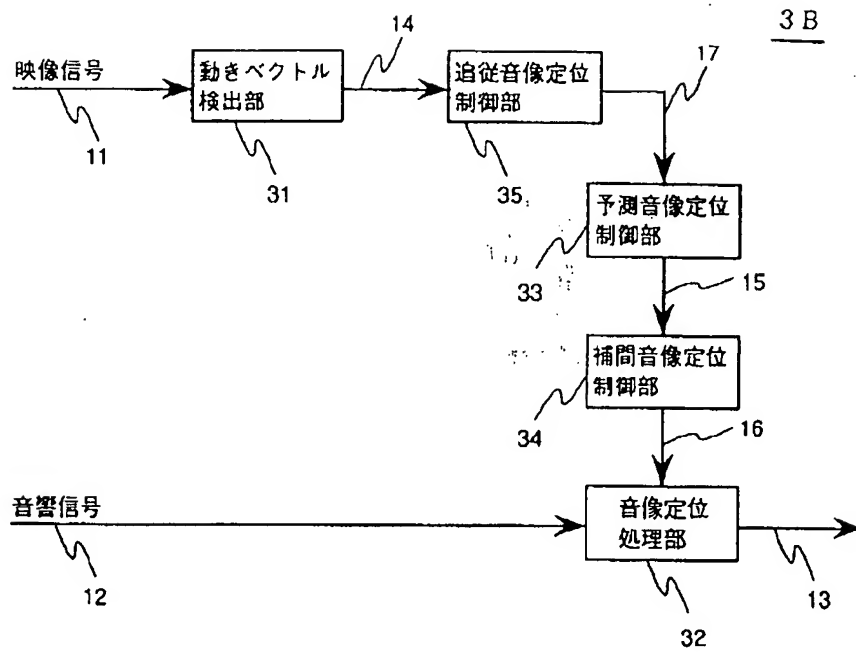


【図3】



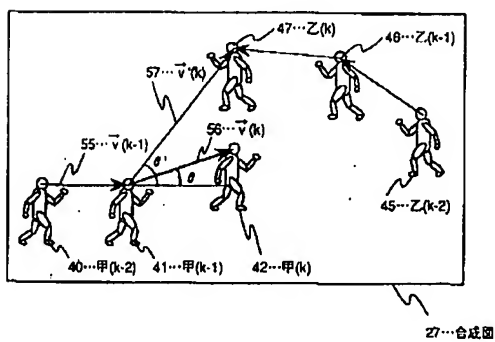
【図5】

(図5)



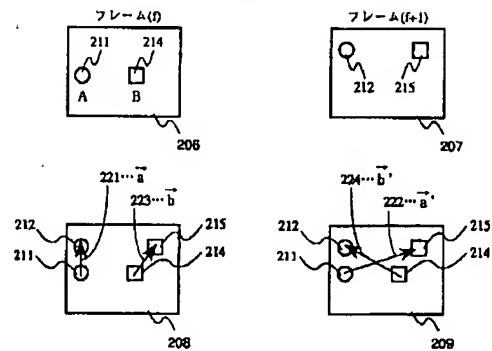
【図6】

(図6)

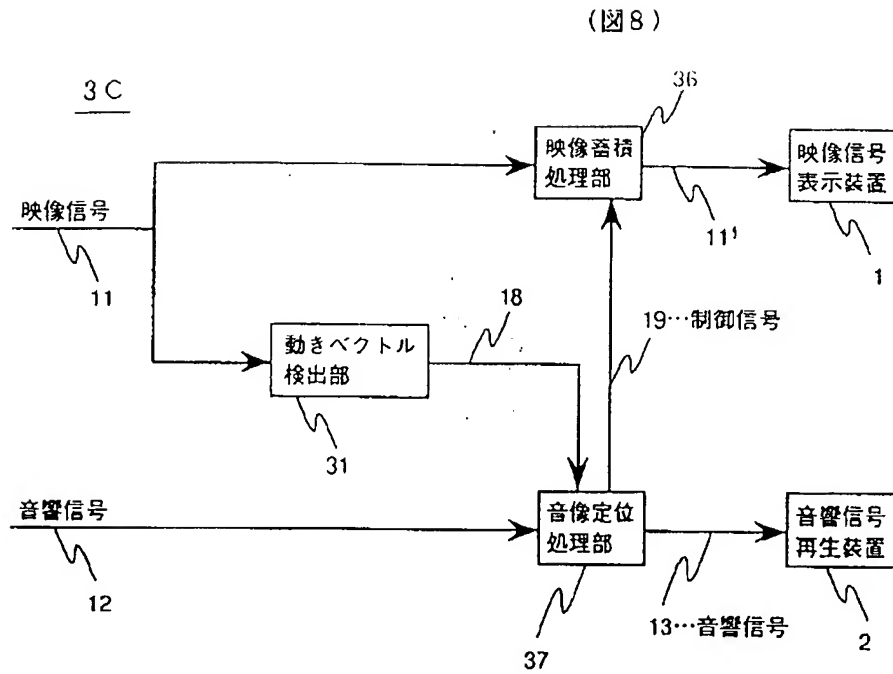


【図7】

(図7)



【図8】



【図9】

